Journal Pre-proof

Distance problems in the behavioral setting

Antonio Fazzi, Ivan Markovsky

PII: S0947-3580(23)00061-4
DOI: https://doi.org/10.1016/j.ejcon.2023.100832
Reference: EJCON 100832

To appear in: *European Journal of Control*

Received date: 3 May 2023
Accepted date: 8 June 2023

Please cite this article as: Antonio Fazzi, Ivan Markovsky, Distance problems in the behavioral setting, *European Journal of Control* (2023), doi: https://doi.org/10.1016/j.ejcon.2023.100832

European Journal of Control 000 (2023) 1–6

Contents lists available at ScienceDirect

European Journal of Control

journal homepage: www.elsevier.com/locate/ejcon

Distance problems in the behavioral setting ☆

Antonio Fazzi [a,*], Ivan Markovsky [b,c]

[a] Department of Information Engineering, University of Padova, I-35131 Padova, Italy
[b] International Centre for Numerical Methods in Engineering (CIMNE), Gran Capitn, 08034 Barcelona, Spain
[c] Catalan Institution for Research and Advanced Studies (ICREA), Pg. Lluis Companys 23, Barcelona, Spain

ARTICLE INFO

Article history:
Received 3 May 2023
Accepted 8 June 2023
Available online xxx

Keywords:
Computational methods
linear systems
Distance

ABSTRACT

Motivated by the *distance to uncontrollability* problem, we define a distance between finite-length linear time-invariant behaviors. The method proposed in this paper for computing the distance exploits the principal angles associated with structured matrices representing the systems.

1. Introduction

The concept of distance between linear time-invariant (LTI) systems is important in systems and control theory. For example, it is used to measure the robustness of a given system with respect to a property, e.g., controllability or stability. In the input-output setting, the *gap metric* [1–3] defines a distance between LTI systems. An equivalent metric (i.e., inducing the same topology) is the *graph metric* [4,5]. The computation of the gap metric [6] was shown to be equivalent to the solution of an optimization problem involving (right) coprime factorizations of the systems. Another metric called the Vinnicombe metric, or the ν-gap [7,8], is used for stability analysis.

It can provide a more stringent test for robustness with respect to the gap metric and it is easier to compute since it involves the computation of the winding number of a rational function.

Both the gap metric and the Vinnicombe metric are implemented in the Matlab function *gapmetric* [9]. More distances associated with input-output maps are the $\mathcal{L}_2$ gap, which is cheap to compute since it involves only a norm computation but is not useful in the context of robust stability [8], and the Sasane-Ball metric [10] which extends the notion of *gap* to linear systems having nonzero initial conditions.

The work by Ball and Sasane connects the classical gap metric with the *behavioral approach* to system theory [11,12] introduced by J. C. Willems. In the behavioral framework, LTI systems are defined as sets of trajectories, without distinction between input and output variables. The behavioral approach motivated the study of the connection between the classical gap metric defined for input-output maps and the distance between behaviors [13,14,10,15,16]. In particular, [15] extends four of the previously mentioned metrics (the classical gap metric [1], the $\mathcal{L}_2$ metric [8], the Sasane-Ball metric [10] and the Vinnicombe metric [8,7]) from the classical input-output setting to the behavioral setting by using a rational behavioral representation [17,18].

The recent work in [16], on the other hand, proposes a trajectory-based idea of distance for finite length behaviors by showing its connection with the classical concept of *gap*. Its computation is easy and is based on the singular value decomposition of the Hankel matrix built from an observed trajectory. We remark that all these distances are restricted to the subset of controllable behaviors. The controllability assumption is removed in [16] by assuming a rank condition on the Hankel matrix generated by the system trajectory [19]. A definition of distance between behaviors without restrictions is motivated by a model reduction problem in the behavioral setting [20].

The goal of this paper is to define a new metric in the behavioral setting that is intuitive, computationally cheap, and applies to an arbitrary LTI system (unstable, uncontrollable, *etc.*). Such a metric is useful in distance problems, *e.g.*, the distance to uncontrollability, where systems representations are used in the definition of distance (see [21] for details). Therefore, in this paper, we define a distance independent of the choice of the system representation. Inspired by the previous works, we propose an adaptation of the concept of principal angles, and we define different distance measures as functions of the principal angles.

☆ This work was not supported by any organization
* Corresponding author.
E-mail addresses: antonio.fazzi@unipd.it (A. Fazzi), imarkovsky@cimne.upc.edu (I. Markovsky).

https://doi.org/10.1016/j.ejcon.2023.100832

2. Preliminaries and problem statement

2.1. Behavioral system theory

In the behavioral setting [11], LTI systems are defined as sets of trajectories. This allows for stating and defining system properties in terms of the observed trajectories of the system.

An LTI dynamical system with q variables, $\mathcal{B}^q$, is characterized by a set of integer invariants [22]: q is the number of variables, m is the number of inputs, p is the number of outputs, ℓ the lag, and n the order. These integers are a property of the system and do not depend on the choice of representation. The complexity of the system is defined as the pair (m, n) — (number of inputs, order).

We denote by $\mathcal{B}|_L$ the behavior $\mathcal{B}$ restricted to the interval $[1, L]$, that is we truncate all trajectories to time L. By restricting the behavior to time L, we can write its dimension (as a vector space) in terms of the complexity [23]:

$$dim\ \mathcal{B}|_L = n + mL.$$

2.2. Structured matrices

Given a linear time-invariant system, there are different ways of representing its behavior. The notation $w \in \mathcal{B}|_L$ means that w is a L-samples long trajectory of the system. A length-T trajectory w satisfies a difference equation:

$$\mathcal{B}(R) = \{w | R_0 w(t) + R_1 \sigma w(t) + \cdots + R_l \sigma^l w(t) = 0\}, \qquad (1)$$

for all $t = 1, \ldots, T - l$, where $R(\sigma) \in \mathbb{R}^{p \times q}[\sigma]$ is a matrix polynomial of degree ℓ, usually called kernel representation, and σ is the shift operator $\sigma w(t) = w(t + 1)$.

By writing (1) in extended form as

$$(R_0, R_1, \cdots, R_\ell) \begin{pmatrix} w(1) & w(2) & \cdots & \iddots \\ w(2) & w(3) & \cdots & \\ \vdots & \vdots & \iddots & \\ w(\ell+1) & w(\ell+2) & & w(T) \end{pmatrix} = 0, \qquad (2)$$

we observe the arising of a Hankel matrix $H_{\ell+1}(w)$ (the subscript denotes the number of rows) generated by the entries of the trajectory w. A classical result in the behavioral setting, also known as *fundamental lemma* [24], states that, under suitable assumptions, the finite-length behavior $\mathcal{B}_L$ can be generated by one observed trajectory only by considering linear combinations of the columns of the Hankel matrix $H_{\ell+1}(w)$:

Lemma 1. *Given a behavior $\mathcal{B}^q$ of order n, let $w \in \mathcal{B}^q_T$. Assume $w = (u, y)$ is an input-output partition of the variables. If*

1. *$\mathcal{B}^q$ is controllable,*
2. *the input component u is persistently exciting of order $L + n$[1], for a certain $1 \leq L \leq T$,*

then $\mathcal{B}^q_L$ = image $H_L(w)$

It has been recently shown in [19] that the assumptions in Lemma 1 can be replaced by a rank condition on the Hankel matrix $H_L(w)$: this is important since it allows extending the same result to systems that are not necessarily controllable. However, a trajectory-based definition of distance has been already proposed in [16], so we plan to use a different structured matrix associated with the full behavior $\mathcal{B}^q_L$.

By writing (2) in a different way, we get the following result from [25]:

[1] u is persistently exciting of order L if the Hankel matrix $H_L(u)$ is full row rank.

Lemma 2. *Given a behavior $\mathcal{B}_L$ expressed by its kernel representation $\mathcal{B}_L = ker\ R(\sigma)$, let $R^1(\sigma), \ldots, R^p(\sigma)$ be the rows of $R(\sigma)$. Then we have*

$$\mathcal{B}|_L = ker\ \mathcal{T}_L(R) = ker \begin{bmatrix} \mathcal{T}_L(R^1) \\ \vdots \\ \mathcal{T}_L(R^p) \end{bmatrix}, \quad for\ L \geq \ell + 1,$$

where ℓ is the degree of $R(\sigma)$ and $\mathcal{T}_L(R^1), \ldots, \mathcal{T}_L(R^p)$ are generalized Toeplitz matrices with L block columns

$$\mathcal{T}_L(R^i) = \begin{bmatrix} R^i_0 & R^i_1 & \cdots & R^i_\ell & & \\ & R^i_0 & R^i_1 & \cdots & R^i_\ell & \\ & & \ddots & \ddots & & \ddots \\ & & & R^i_0 & R^i_1 & \cdots & R^i_\ell \end{bmatrix}, \qquad (3)$$

for $i = 1, \ldots, p$.

The previous result holds true for any linear time-invariant system. Moreover, the kernel representation does not change with the length of the trajectory, even if the Toeplitz matrix $T_L(R)$ does. But the kernel representation $R(\sigma)$ is not unique. The goal is to define a distance measure that does not depend on the particular choice for the kernel representation.

2.3. Problem statement

As already stated, the kernel representation $R(\sigma)$ is not unique, but it is well known that an equivalent kernel representation can be obtained by pre-multiplication by a unimodular matrix $U(\sigma)$, that is a square matrix polynomial whose determinant is a nonzero constant. First of all, we will work with minimal kernel representations.

Remark 3. *Kernel representations are matrix polynomials having q columns and (at least) p rows. If the number of rows is exactly p, the kernel representation is minimal, that is its rows are linearly independent on the ring of (scalar) polynomials. Every kernel representation can be reduced to a minimal one by applying suitable transformations on its rows, therefore we can assume all the kernel representations to be minimal without loss of generality.*

We observe that the pre-multiplication by an arbitrary unimodular matrix $U(\sigma)$ can, in general, change the degree of $R(\sigma)$, that is $deg(U(\sigma)R(\sigma)) > deg(R(\sigma))$. However, by considering Hankel matrices with $\ell + 1$ rows (we recall that the lag ℓ is an invariant integer for the system) and their left kernel(s), we can also assume that all the kernel representations have the same (minimum) degree ℓ.

Remark 4. *By constraining the degree of all the kernel representations to be minimum, we need to restrict the class of equivalent representations. The unimodular matrix polynomial has, in this case, a special form. It is a (square) invertible matrix whose entries are scalar.*

The goal is to define a distance between two behaviors $\mathcal{B}^1_L(R^1)$ and $\mathcal{B}^2_L(R^2)$ which does not change by premultiplying R^1 and R^2 by invertible matrices (see Remark 4). More formally:

Problem 5. *Given two finite-length behaviors $\mathcal{B}^1_L(R^1), \mathcal{B}^2_L(R^2)$, define a distance which satisfies the following property:*

$$\begin{aligned} &dist(\mathcal{B}^1_L(R^1), \mathcal{B}^2_L(R^2)) = \\ &dist(\mathcal{B}^1_L(\Theta_1 R^1), \mathcal{B}^2_L(\Theta_2 R^2)) \quad \forall \Theta_1, \Theta_2 \in \mathbb{R}^{p \times p}. \end{aligned} \qquad (4)$$

3. Motivating application

The distance to uncontrollability problem in the behavioral setting has been already studied and analyzed in [26] for SISO systems and in [21] for MIMO systems. This problem was stated as a distance

to singularity for the kernel representation $R(\sigma)$. Once we split the set of variables into inputs and outputs, $w = (u, y)$, the key idea is to partition accordingly the kernel representation as $R(\sigma) = [P(\sigma), Q(\sigma)]$. This distance problem was shown, then, to be equivalent to a distance to common divisibility for the two (matrix) polynomials $P(\sigma), Q(\sigma)$ appearing in this input-output representation. And it was solved numerically with the gradient system methodology proposed in [27,28] for scalar polynomials and in [29] for matrix polynomials. The considered polynomial distance was the classical norm of the difference between the polynomial coefficients.

But the distance between polynomial coefficients is representation invariant in the SISO case only. When we deal with MIMO systems, we need to fix a particular kernel representation to get a well-posed definition of distance. To solve this issue, we are going to propose an alternative definition of distance which still depends on the kernel representations of the systems, but it has the same value for each equivalent kernel representation:

$$dist(R^1(\sigma), R^2(\sigma)) = dist(\theta_1 R^1(\sigma), \theta_2 R^2(\sigma)) \ \forall \theta_1, \theta_2 \in \mathbb{R}^{p\times p}$$

We observe that controllability is a classical example of a property that is usually defined in terms of system trajectories (see, e.g., [23, Section 7.5]), but its test relies on the system kernel representation. There are no trajectory-based characterizations of distance to uncontrollability, up to our knowledge.

Remark 6. *There are differences in the distance measures to uncontrollability between the SISO [26, Section 3] and the MIMO [21, Section II] cases. Distance measures for SISO systems are representation invariant (see Section 4.1), while the same is not true in the MIMO case. This motivates the introduction of the proposed distance.*

4. Principal angles between behaviors

The principal angles are an extension of the classical angle between two lines, and they are some quantities that are invariant under isometric transformations. The distance we are going to propose is based on the principal angles between the subspaces generated by the Toeplitz matrices representing the two behaviors, as shown in Lemma 2. The choice of the principal angles is motivated by the following

Theorem 7. *[30, Theorem 2] Any notion of distance between k-dimensional subspaces in $\mathbb{R}^n$ that depends only on the relative positions of the subspaces, i.e., it is invariant under any rotation, must be a function of their principal angles.*

Moreover, by looking at [31, Section 4.3], several distance measures in terms of the principal angles are proposed, so it is also possible to choose the one which best fits the considered problem. Because of Lemma 1 and the rank condition on the Hankel matrix [19], behaviors can be seen as low-order subspaces in a bigger vector space, that is as points on a Grassmann manifold [32]. Therefore, suitable choices could be the arc-length distance $dist = \|\theta\|_2$ (this is the geodesic distance [33]) or the classical projection distance [34] (which is the choice adopted in [16]).

The computation of such angles between the two Toeplitz matrices follows the algorithm in [34, Chapter 12]. This is based on two simple steps: the (tiny) QR factorization of the two matrices and the computation of a singular value decomposition. We list the main steps for the computation of the angles and the associated distance in the following:

1. Compute the tiny QR decomposition of the two matrices
$$\begin{aligned} T_{L_1}(R_1)^T &= Q_1 T_1 \\ T_{L_2}(R_2)^T &= Q_2 T_2 \end{aligned}$$
2. Compute the singular values Σ of the matrix $Q_1^T Q_2$
3. $\theta = \arccos \Sigma$
4. Compute the distance as a function of θ.

To fix the ideas, we define $dist(R_1, R_2) = \|\theta\|_2$, but other functions of the principal angles can be fair alternative distances (see [31, Section 4.3]).

Remark 8. *The considered matrices need to have more rows than columns to get a non-trivial solution (not all the singular values of the orthogonal matrix, that is the cosines of the angles, equal to 1). This is why we transposed the Toeplitz matrices in the first step. For the same reason, we need to consider a tiny QR decomposition instead of a full one.*

The Toeplitz matrices and their QR decompositions actually still depend on the particular kernel representation. The problem of defining a representation invariant measure of distance (according to the previous angle-based definition) is equivalent to the fact that the matrix Q in the QR decomposition does not change by replacing R with ΘR in the kernel representation. We show how to make it possible in the following sections, by analyzing separately the SISO and the MIMO case.

4.1. The SISO case

In the SISO case, a kernel representation of a system has the general form $R(z) = [q(z), p(z)]$, where $p(z), q(z)$ are two scalar polynomials. The number of outputs p equals 1, therefore Θ_1 and Θ_2 in (4) are scalars. Given two SISO behaviors $\mathcal{B}_1(R_1), \mathcal{B}_2(R_2)$ having the same order, we consider two different equivalent kernel representations for each of them,

$$\begin{aligned} R_i(z) &= [q_i(z), p_i(z)] \quad i = 1, 2 \\ \Theta_i R_i(z) &= [\Theta_i q_i(z), \Theta_i p_i(z)] \quad \Theta_i \in \mathbb{R}, \ i = 1, 2. \end{aligned} \tag{5}$$

The problem is to understand when $dist(R_1(z), R_2(z)) = dist(\Theta_1 R_1(z), \Theta_2 R_2(z)) \quad \forall \Theta_1, \Theta_2$.

We need the following result:

Lemma 9. *Given a matrix A and its QR factorization $A = QT$, multiplying A (on the right) by an upper triangular matrix S does not change the matrix Q in the QR factorization.*

Proof. $AS = QTS = Q(TS)$. Since S is upper triangular, the matrix TS is the upper triangular factor in the QR decomposition of AS. □

The multiplication of all the coefficients in kernel representation by Θ changes the Toeplitz matrix from $\mathcal{T}_L(R)$ to $\mathcal{T}_L(\Theta R)$ such that

$$\begin{aligned} T_L(\Theta R) &= diag(\Theta) T_L(R) \Rightarrow \\ T_L(\Theta R)^T &= T_L(R)^T diag(\Theta)^T = T_L(R)^T diag(\Theta) \end{aligned} \tag{6}$$

Remark 10. *Since Θ, as well as all the coefficients, are scalar, in the SISO case we do not care about left and right multiplications.*

By combining Lemma 9 and (6), the matrix Q does not change in the QR factorization by changing the kernel representation. That is, in the SISO case, the choices for the distance measure (depending on the principal angles) and the behaviors representations automatically give a representation invariant distance.

4.2. The MIMO case

In the MIMO case, a kernel representation $R(\sigma) = [Q(\sigma), P(\sigma)]$ is, in general, a matrix polynomial (a polynomial with matrix coefficients). Changing kernel representation, according to Remark 4,

means pre-multiplying by a $p \times p$ invertible matrix. We wonder if the same results of Section 4.1 hold true also in this case, but the answer is negative, in general.

Given a matrix polynomial $R(\sigma)$, we observe that its pre-multiplication by Θ changes the corresponding transposed Toeplitz matrix as follows:

$$\begin{aligned} \mathcal{T}_L(\Theta R)^T &= \begin{pmatrix} R_0^T\Theta^T & & \\ R_1^T\Theta^T & R_0^T\Theta^T & \\ \vdots & \ddots & \\ & & R_\ell^T\Theta^T \end{pmatrix} = \\ &= \begin{pmatrix} \bar{R}_0^T & & \\ \bar{R}_1^T & \bar{R}_0^T & \\ \vdots & \ddots & \\ & & \bar{R}_\ell^T \end{pmatrix} \begin{pmatrix} \Theta^T & & \\ & \ddots & \\ & & \Theta^T \end{pmatrix} = \\ &= \mathcal{T}_L(R)^T \qquad\qquad diag(\Theta)^T \end{aligned} \tag{7}$$

The key point is now the structure of the matrix $diag(\Theta^T)$: if Θ^T is upper triangular, $diag(\Theta^T)$ is upper triangular too, and the same result from Section 4.1 (based on Lemma 9) still holds true. But this is not general, though. If Θ^T is not upper triangular, Lemma 9 cannot be applied and the orthogonal matrix in the QR decomposition changes by switching kernel representation.

To deal with this issue, we can use a *block* QR decomposition which takes into account the dimension of the matrix Θ and computes a block-upper triangular matrix in the QR decomposition. This means that the diagonal coefficients of such a matrix are $p \times p$ blocks (of the same dimension of Θ^T). Doing so, we expect similar invariance properties as the ones observed in Section 4.1. But such decomposition is not available in standard software packages, so its computation needs to be implemented in an algorithm.

The idea for this block factorization comes from the extension of the classical Householder method [35–37]. We summarize in Algorithm 1 the code from [35].

Algorithm 1: Block QR decomposition

Data: A (matrix to be decomposed), r (size of square blocks on the diagonal of rectangular factor)
Result: Q (orthogonal), T (block upper triangular) such that $A = QT$
begin
- Get the number of block columns ℓ
- **for** $i = 1 : \ell$ **do**
  - Let C be the i-th block column of A
  - Compute $C = XZ$ (tiny QR factorization)
  - Set as $\hat{X}$ the first r rows of X
  - Compute the SVD $\hat{X} = WDV^T$
  - Set $Y = X + \begin{pmatrix} WV^T \\ 0 \end{pmatrix}$
  - Compute $Y = US$ (tiny QR factorization)
  - Set $Q_i = I - 2UU^T$, $T_i = Q_i * A$, $\hat{Q}_i = \begin{pmatrix} I & 0 \\ 0 & Q_i \end{pmatrix}$
  - Delete from A the first r columns and the first r rows
- **for** $j = 1 : \ell$ **do**
  - $Q = Q * \hat{Q}_j$
- $Q = Q(:, 1 : r\ell), T = T(1 : r\ell, 1 : r\ell)$

By applying Algorithm 1, the orthogonal factor Q still preserves its properties (in the classical sense), but the structure of the triangular factor T is a bit different from the classical one, since its diagonal elements are $r \times r$ matrices (see Figure 1 for an example with $r = 2$).

Fig. 1. Illustrative example of a block upper triangular factor in a block QR decomposition. The dimension of the blocks is 2×2.

We can state now the following result:

Lemma 11. *Given a matrix A and its block QR factorization $A = QT$, multiplying A (on the right) by a block upper triangular matrix does not change the orthogonal factor Q in the block QR factorization.*

Proof. It can be checked that multiplying two block upper triangular matrices gives a matrix with the same structure. Hence, if S is a block upper triangular matrix, $AS = (QT)S = Q(TS)$. Since T is block upper triangular, the matrix TS is the block upper triangular factor in the block QR decomposition of AS. □

To conclude, by applying this block factorization, we leave unchanged the orthogonal matrix Q in the QR factorization. The computation of the principal angles, the choice for the behavior representation and Lemma 11 allow having a representation invariant measure of distance.

5. Numerical examples

We show some numerical examples with MIMO systems to illustrate the previous results.

We consider first the distance between a controllable system and the closest uncontrollable one. This problem was stated and solved numerically by a local optimization approach in [21]. The problem in [21] was solved as an approximate polynomials common factor computation, hence the considered (representation-based) distance depends on the coefficients of the polynomials. The proposed method, on the other hand, is invariant on the choice of the representation (despite its computation requiring a representation).

Therefore, differently from [16], both these distances are representation based. The classical definition of controllability in the behavioral setting is the possibility of linking any two system trajectories, up to a delay of time. It can be checked by a rank-test on the system representation. Therefore, a distance measure based on the representations is advisable.

We consider a MIMO system with two inputs and two outputs, by generating two matrix polynomials (of degree 2) which form its kernel representation $[Q_c(z), P_c(z)] = R_c(z)$. We look for the closest uncontrollable system $R_u(z) = [Q_u(z), P_u(z)]$ by an *approximate common factor* between $Q_c(z)$ and $P_c(z)$ (see [21] for the details). We compute then the two distances

1. distance between polynomials coefficients: $d_{PQ} = \sqrt{\|P_c - P_u\|_F^2 + \|Q_c - Q_u\|_F^2}$
2. distance defined by principal angles (see Section 4): $dist(R_c, R_u) = \|\theta\|_2$

$$\begin{aligned} Q_c(z) &= \begin{pmatrix} -2 & 0 \\ 0 & 2 \end{pmatrix} z^2 + \begin{pmatrix} 2 & -3 \\ 3 & 0 \end{pmatrix} z + \begin{pmatrix} -5 & 0 \\ -3 & 1 \end{pmatrix} \\ P_c(z) &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} z^2 + \begin{pmatrix} -2 & -1 \\ 5 & 1 \end{pmatrix} z + \begin{pmatrix} 0 & -2 \\ 2 & -5 \end{pmatrix} \end{aligned} \tag{8}$$

It can be checked that $P_c(z)$ and $Q_c(z)$ have no common roots, hence the system associated with $R_c = [Q_c, P_c]$ is controllable. The representation of the closest uncontrollable system (computed by the

algorithm in [29]) is (the following coefficients are rounded to the third decimal digit)

$$
\begin{aligned}
Q_u(z) = &\begin{pmatrix} -1.276 & -0.536 \\ 0.657 & 1.398 \end{pmatrix} z^2 + \begin{pmatrix} 2.125 & -2.973 \\ 3.206 & -0.142 \end{pmatrix} z + \\
&\begin{pmatrix} -5.078 & 0.007 \\ -2.990 & 1.042 \end{pmatrix} \\
P_u(z) = &\begin{pmatrix} 0.462 & 0.280 \\ -0.175 & 1.207 \end{pmatrix} z^2 + \begin{pmatrix} -2.089 & -0.843 \\ 4.875 & 1.117 \end{pmatrix} z + \\
&\begin{pmatrix} -0.010 & -2.019 \\ 1.954 & -5.042 \end{pmatrix}
\end{aligned} \tag{9}
$$

We have $d_{PQ} = 1.483$ and $dist(R_c, R_u) = 0.598$. What happens by switching to equivalent representations for both (8) and (9)? Consider

$$
\begin{aligned}
\bar{R}_c = &\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} R_c = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} [Q_c, P_c] = [\bar{Q}_c, \bar{P}_c] \\
\bar{R}_u = &\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} R_u = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} [Q_u, P_u] = [\bar{Q}_u, \bar{P}_u]
\end{aligned} \tag{10}
$$

$\bar{R}_c$ and $\bar{R}_u$ represent the same systems (the same trajectories satisfy the associated difference equations). Hence, it would be nice if the computed value of distance does not change with the representation. However, this is true only for the distance proposed in the paper because the polynomial coefficients are completely different. Indeed we have $d_{\bar{P}\bar{Q}} = 19.623$ and $dist(\bar{R}_c, \bar{R}_u) = 0.598$.

Remark 12. *The value $d_{\bar{P}\bar{Q}}$ only shows that the value of the distance between the coefficients changes with the system representation. It is not the distance of $\bar{R}_c$ to the closest uncontrollable system. Despite R_c and $\bar{R}_c$ representing the same system, the algorithm in [29] optimizes with respect to the starting polynomials (hence the need to fix a representation). Equivalent representations lead to different solutions.*

The invariance property of the proposed distance can be further checked on equivalent representations of random systems, generated as follows:

```
q = 4;
m1 = 2;
m2 = 2;
l1 = 1;
l2 = 1;
T = 100;

p1 = q - m1;
p2 = q - m2;
n1 = l1 * p1;
n2 = l2 * p2;

sys1 = drss(n1, p1, m1);
sys2 = drss(n2, p2, m2);

u1 = randn(T, m1);
y1 = lsim(sys1, u1);
w1 = [u1 y1];
u2 = randn(T, m2);
y2 = lsim(sys2, u2);
w2 = [u2 y2];

R1 = null(blkhank(w1, l1 + 1, T - l1)')'
R2 = null(blkhank(w2, l2 + 1, T - l2)')'

P1 = randn(p1, p1);
R3 = P1 * R1;
P2 = randn(p2, p2);
R4 = P2 * R2;
```

By using the factorization in Algorithm 1, the principal angles give $dist(R1, R2) = dist(R3, R4) = dist(R1, R4) = dist(R3, R2)$. This is because the principal angles do not change by switching between equivalent representations.

6. Conclusion

Working in the behavioral setting, where linear time-invariant systems are defined as sets of trajectories, we proposed a new definition of distance measure based on the principal angles associated with some structured matrices representing finite-length behaviors. The proposed definition of distance is based on the system kernel representation; such a representation is nonunique, but a suitable matrix factorization allows to return the same distance measure for each equivalent representation.

We remark that we only defined a distance measure. Such a measure can be used to develop algorithms that compute the numerical solution of some (representation invariant) distance problems.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The research leading to these results has received funding from: the Catalan Institution for Research and Advanced Studies (ICREA), the Fond for Scientific Research Vlaanderen (FWO) projects G090117N and G033822N; and the Fonds de la Recherche Scientifique FNRS–FWO EOS Project 30468160.

This project was partially supported by the Italian Ministry of University and Research under the PRIN'17 project "Data-driven learning of constrained control systems", contract no. 2017J89ARP.

References

[1] G. Zames, A. El-Sakkary, Uncertainty in unstable systems: The gap metric, IFAC Proceedings Volumes 14(2) (1981) 149–152. 8th IFAC World Congress on Control Science and Technology for the Progress of Society, Kyoto, Japan, 24-28 August 1981. 10.1016/S1474-6670(17)63475-X

[2] A. K. El-Sakkary, The gap metric: Robustness of stabilization of feedback systems, IEEE Trans. Automat. Control AC-30 (1985) 240–247.

[3] T. T. Georgiou, M. C. Smith, Optimal robustness in the gap metric, IEEE Trans. Automat. Contr. 35 (1990) 673–686.

[4] M. Vidyasagar, The graph metric for unstable plants and robustness estimates for feedback stability, IEEE Trans. Automat. Control AC-29 (1984) 403–418.

[5] M. Vidyasagar, H. Schneider, B. A. Francis, Algebraic and topological aspects of feedback stabilization, IEEE Trans. Automat. Control 27 (1982) 880–894.

[6] T. T. Georgiou, On the computation of the gap metric, Syst. Control Lett. 11 (1988) 253–257.

[7] G. Vinnicombe, Frequency domain uncertainty and the graph topology, IEEE Trans. Automat. Control 38 (1993) 1371–1383.

[8] G. Vinnicombe, Uncertainty and Feedback: H_∞ Loop-Shaping and the ν-Gap Metric, Imperial College Press, London, 2001.

[9] gapmetric, (https://www.it.mathworks.com/help/robust/ref/lti.gapmetric.html). Accessed: 2022-07-16.

[10] J. A. Ball, A. J. Sasane, Equivalence of behavioral distance and the gap metric, Systems Control Lett. 55 (2006) 214–222.

[11] J. W. Polderman, J. C. Willems, Introduction to Mathematical Systems Theory, volume 26 of Texts in Applied Mathematics, Springer New York, New York, NY, 1998. 10.1007/978-1-4757-2953-5

[12] J. C. Willems, The behavioral approach to open and interconnected systems: Modeling by tearing, zooming, and linking, Control Systems Magazine 27 (2007) 46–99.

[13] A. J. Sasane, Distance between behaviours, Int. J. Control 76 (2003) 1214–1223.

[14] W. Bian, M. French, H. K. Pillai, An intrinsic behavioral approach to the gap metric, SIAM J. Control Optim. 47 (2008) 1939–1960.

[15] H. L. Trentelman, S. V. Gottimukkala, Distance between behaviors and rational representations, SIAM J. Control Optim. 51 (2013) 4211–4241.

[16] A. Padoan, J. Coulson, H. J. van Waarde, J. Lygeros, F. Dörfler, Behavioral uncertainty quantification for data-driven control, in: Proc. of the IEEE Conf. on Decision and Control, Cancun, Mexico, 2022. 10.1109/CDC51059.2022.9993002

[17] J. C. Willems, Y. Yamamoto, Behaviors defined by rational functions, Linear Algebra Appl. 425 (2007) 226–241.

[18] S. Gottimukkala, S. Fiaz, H. L. Trentelman, Equivalence of rational representations of behaviors, System and control letters 60 (2011) 119–127.

[19] I. Markovsky, F. Dörfler, Identifiability in the behavioral setting, IEEE Trans. Automat. Contr. (2023). 10.1109/TAC.2022.3209954

[20] B. Roorda, S. Weiland, Optimal angle reduction - a behavioral approach to linear system approximation, Linear Algebra Appl. 337 (2001) 189–235.

[21] A. Fazzi, N. Guglielmi, I. Markovsky, Computing common factors of matrix polynomials with applications in system and control theory, in: Proc. of the IEEE Conf. on Decision and Control, Nice, France, 2019, pp. 7721–7726. 10.1109/CDC40024.2019.9030137

[22] J. C. Willems, From time series to linear system-Part I. Finite dimensional linear time invariant systems, Automatica 22 (1986) 561–580.

[23] I. Markovsky, J. C. Willems, S. Van Huffel, B. De Moor, Exact and Approximate Modeling of Linear Systems: A Behavioral Approach, SIAM, 2006. 10.1137/1.9780898718263

[24] J. C. Willems, P. Rapisarda, I. Markovsky, B. De Moor, A note on persistency of excitation, Syst. Control Lett. 54(4) (2005) 325–329.

[25] A. Fazzi, I. Markovsky, Addition and intersection of linear time-invariant behaviors, 2022. https://arxiv.org/abs/2107.01588.

[26] I. Markovsky, A. Fazzi, N. Guglielmi, Applications of polynomial common factor computation in signal processing, in: Y. D. et al. (Ed.), Latent Variable Analysis and Signal Separation, volume 10891 of Lecture Notes in Computer Science, Springer, 2018, pp. 99–106. 10.1007/978-3-319-93764-9_10

[27] N. Guglielmi, I. Markovsky, An ODE based method for computing the distance of co-prime polynomials to common divisibility, SIAM Journal on Numerical Analysis 55 (2017) 1456–1482. 10.1137/15M1018265

[28] A. Fazzi, N. Guglielmi, I. Markovsky, An ODE based method for computing the Approximate Greatest Comnion Divisor of polynomials, Numer. Algorithms 81 (2019) 719–740. 10.1007/s11075-018-0569-0

[29] A. Fazzi, N. Guglielmi, I. Markovsky, Generalized algorithms for the approximate matrix polynomial GCD of reducing data uncertainties with application to MIMO system and control, J. Comput. Appl. Math. 393 (2021). 10.1016/j.cam.2021.113499

[30] K. Ye, L. H. Lim, Schubert varieties and distances between subspaces of different dimensions, SIAM J. Mat. Anal. Appl. 37 (2016) 1176–1197.

[31] A. Edelman, T. A. Arias, S. T. Smith, The geometry of algorithms with orthogonality constraints, SIAM J. Matrix Anal. Appl. 20 (1998) 303–353.

[32] K. Usevich, I. Markovsky, Optimization on a Grassmann manifold with application to system identification, Automatica 50 (2014) 1656–1662. 10.1016/j.automatica.2014.04.010

[33] X. Wang, Z. Li, D. Tao, Subspaces indexing model on grassmann manifold for image search, IEEE Trans. Image Process 20 (2011) 2627–2635.

[34] G. Golub, V. L. C., Matrix Computation, Johns Hopkins University Press, Baltimore, Maryland, 3rd edition, 1996.

[35] H. Murakami, An implementation of the block householder method, IPSJ Digital Courier 2 (2006) 298–317. 10.2197/ipsjdc.2.298

[36] F. Rotella, I. Zambettakis, Block householder transformation for parallel qr factorization, Appl. Math. Lett. 12 (1999) 29–34.

[37] R. Schreiber, B. Parlett, Block reflectors: Theory and computation, SIAM J. Numer. Anal. 25(1) (1988) 189-205. 10.1137/0725014. 10.1137/0725014